
Berliner Debatte Initial

4

27. Jg. 2016

Big Data als Theorieersatz

Broemel,
Trute

Alles nur Datenschutz?

Mainzer

Veränderungen des Theoriebegriffs

Böschchen,
Huber,
König

Algorithmische Subpolitik

Sacks

Die „weiße Suche“ nach Joe Louis

Mierau

elektronische Sonderausgabe
der Druckfassung mit
ISBN 978-3-945878-11-8
www.berlinerdebatte.de

Sergej Tretjakow und die DDR

Autorinnen und Autoren

Alexander Amberger, Dr.,
Politikwissenschaftler, „Helle Panke“ e.V. –
Rosa-Luxemburg-Stiftung Berlin

Stefan Böschen, PD Dr., Dipl.-Ing.,
Chemieingenieur und Soziologe, Leiter des
Forschungsbereichs Wissensgesellschaft und
Wissenspolitik, Institut für Technikfolgenab-
schätzung und Systemanalyse am Institut für
Technologie (KIT), Karlsruhe

Roland Broemel, Jun.-Prof. Dr.,
Rechtswissenschaftler, Universität Hamburg

Ulrich Busch, Doz., Dr. habil.,
Finanzwissenschaftler, Leibniz-Sozietät der
Wissenschaften zu Berlin

Wladislaw Hedeler, Dr.,
Historiker, Berlin

Cornelia Heintze, Dr.,
Politikwissenschaftlerin, Leipzig

Georg Huber, M.A., LL.M.,
Rechtswissenschaftler, KIT, Karlsruhe

René König, Dipl.-Soz.,
Soziologe, KIT, Karlsruhe

Klaus Mainzer, Prof. em. Dr.,
Philosoph und Wissenschaftstheoretiker,
Technische Universität München

Jochen Mayerl, Jun.-Prof. Dr.,
Sozialwissenschaftler, Technische Universi-
tät Kaiserslautern

Fritz Mierau,
Slawist, Übersetzer und Herausgeber, Berlin

Thomas Müller, M.A.,
Erziehungswissenschaftler, Berlin

Gregor Ritschel, M.A.
Politikwissenschaftler, Halle

Marcy S. Sacks, Ph.D.,
Professorin, Historikerin, Department of
History, Albion College, Michigan, USA

Jan-Felix Schrape, Dr.,
Soziologe, Universität Stuttgart

Hans-Heinrich Trute, Prof. Dr.,
Rechtswissenschaftler, Universität Hamburg

Loïc Wacquant, Prof. Dr.,
Soziologe, University of California, Berkeley,
Centre européen de sociologie et de science
politique, Paris

Christian Wadepful, M.A.,
Philosoph und Germanist, KIT Karlsruhe

Katharina Anna Zweig, Prof. Dr.,
Informatikerin, Technische Universität
Kaiserslautern

Big Data als Theorieersatz

Zusammengestellt von Gregor Ritschel und Thomas Müller

Editorial	2	***	
		<i>Marcy S. Sacks</i>	
		Vielsagendes Schweigen?	
BIG DATA ALS THEORIEERSATZ		Die Versuche weißer Amerikaner, Joe Louis für sich zu deuten	84
<i>Gregor Ritschel, Thomas Müller</i>			
Big Data als Theorieersatz?	4	<i>Fritz Mierau</i>	
<i>Jan-Felix Schrape</i>		„Vertraulichkeiten zwischen der Kunst und dem Leben“?	
Big Data: Informatisierung der Gesellschaft 4.0	12	Sergej Tretjakow im intellektuellen Haushalt der DDR	93
<i>Klaus Mainzer</i>		<i>Loïc Wacquant</i>	
Zur Veränderung des Theoriebegriffs im Zeitalter von Big Data und effizienten Algorithmen	22	Eine kurze Genealogie und Anatomie des Habitusbegriffs	103
<i>Christian Wadephul</i>		<i>Cornelia Heintze</i>	
Führt <i>Big Data</i> zur abduktiven Wende in den Wissenschaften?	35	Kapitalismusvarianten, Gender und die gesellschaftliche Bedeutung sozialer Dienste	110
<i>Roland Broemel, Hans-Heinrich Trute</i>		BESPRECHUNGEN UND REZENSIONEN	
Alles nur Datenschutz? Zur rechtlichen Regulierung algorithmen- basierter Wissensgenerierung	50	Karl Renner und Viktor Tschernow	
<i>Stefan Böschen, Georg Huber, René König</i>		Zwei neue Abhandlungen über Leben und Werk undogmatischer Sozialisten	
Algorithmische Subpolitik: Big Data als Technologisierung kollektiver Ordnungsbildung?	66	Besprochen von <i>Wladislaw Hedeler</i>	125
<i>Jochen Mayerl, Katharina Anna Zweig</i>		Jürgen Kremer:	
Digitale Gesellschaft und Big Data: Thesen zur Zukunft der Soziologie	77	Geld ohne Schuld. Geldsysteme und Vollgeldreform Rezensiert von <i>Ulrich Busch</i>	129

Ulrich Mählert (Hg.): Die DDR als Chance. Neue Perspektiven auf ein altes Thema Rezensiert von <i>Alexander Amberger</i>	131	INHALTSVERZEICHNIS 27. JAHRGANG 2016 Systematisches Inhaltsverzeichnis	138
Dirk Braunstein: Adornos Kritik der politischen Ökonomie Rezensiert von <i>Ulrich Busch</i>	135	Alphabetisches Inhaltsverzeichnis	141

Editorial

Donald Trumps Sieg bei den US-Präsidentenschaftswahlen hat eindrucksvoll bewiesen, dass der aufreizend laxer Umgang mit Fakten trotz aller moralischen Entrüstung zum politischen Erfolg führen kann. 2016 wird nicht zuletzt wegen Trump als ein Jahr in Erinnerung bleiben, in dem die Grenzen zwischen Wahrheit und Lüge zusehends verwischten. Schon ist wegen der zahllosen Enten, die in sozialen Netzwerken verbreitet und heute *Fake News* genannt werden, vom „postfaktischen Zeitalter“ die Rede. Das ist sicherlich übertrieben, denn kaum jemand kann ernsthaft behaupten, dass in früheren Zeiten nur die Fakten zählten, es keine Lügen gab und niemand belogen werden wollte. Vielleicht lassen sich die Gemüter ein wenig beruhigen, wenn man sich auf ein Wort besinnt, das der US-amerikanische Philosoph Harry Frankfurt berühmt gemacht hat: Bullshit! „Bullshit“ ist jenes Geschwätz, das man auch dann von sich gibt, wenn man keine Ahnung hat, jene heiße

Luft, die man ohne Rücksicht auf Fakten oder Interesse an der Wahrheit absondert. Frankfurt zufolge sollte man Bullshit nicht mit Lügen verwechseln, denn zu lügen setzt voraus, dass man die Wahrheit kennt oder zu kennen meint. Ein nicht geringer Teil der gemeinen und bizarren Behauptungen, die etwa im Netz zu allen möglichen Themen kursieren, scheint hingegen einer Bullshit-Kultur anzugehören, die sich überhaupt nicht um das schert, was zutrifft. Solche Kulturen der Gleichgültigkeit gibt es nicht erst seit Facebook und Co.; ihre zersetzende Wirkung ist längst bekannt. Doch mit den sozialen Netzwerken spitzt sich die Lage dramatisch zu, weil die Hemmschwelle für die Verbreitung von Bullshit drastisch sinkt.

Der Themenschwerpunkt dieses Heftes lenkt den Blick auf ein Phänomen, das die eben skizzierten Entwicklungen flankiert: *Big Data* ist zum Schlagwort für das Sammeln und Auswerten enormer Datensätze geworden, die aus

den digitalen Aktivitäten der Menschen gefiltert und von Algorithmen auf Muster hin analysiert werden. Mit Big Data verbinden sich ebenso große Erwartungen wie Sorgen. Die einzelnen Beiträge fragen einerseits, welche Bedeutung Big Data für die Sozial- und Geisteswissenschaften hat und was von der steilen These, Big Data trete an die Stelle der Theorie und ersetze diese, zu halten ist. Andererseits erkunden die Texte soziale, rechtliche und politische Implikationen von Big Data und demonstrieren damit, dass sozial- und geisteswissenschaftliche Analysen nicht überflüssig werden, sondern nötig sind, um Big Data zu verstehen. Der Themenschwerpunkt greift eine Diskussion in unserer Zeitschrift auf, die Tara Fenwick und Richard Edwards in ihrem Text über die Auswirkungen digitaler Technologien auf Berufe und Professionen angestoßen haben (siehe Heft 1/2016, S. 6-21). Die folgende Einleitung bietet einen Problemaufriss und stellt die Aufsätze des Themenschwerpunkts vor.

Im allgemeinen Teil dieses Heftes laden wir Sie ein, einen neuen Blick auf ein legendäres Sportereignis zu werfen: Am 22. Juni 1938 besiegte Joe Louis im New Yorker Yankee Stadium Max Schmeling in 124 Sekunden im Kampf um die Weltmeisterschaft im Schwergewichtsböxen. Damit nahm er auf überzeugende Weise

Revanche für seine am gleichen Ort erlittene K.o.-Niederlage vom 19. Juni 1936. Sein Sieg war zugleich ein Symbol für die Bereitschaft der USA, sich den Weltmächtsansprüchen Nazi-Deutschlands entschlossen zu widersetzen. Die amerikanische und die deutsche Erinnerung an diese zwei Kämpfe und ihre Protagonisten unterscheidet sich bis heute beträchtlich. Das zeigt sich nicht allein in der abweichenden Gewichtung der zwei Begegnungen, sondern – den verschiedenen, wenn auch zusammenhängenden Sachgeschichten der Länder entsprechend – auch in der Bewertung der beiden Boxer. Mit der Übersetzung eines Aufsatzes von Marcy S. Sacks stellen wir eine inneramerikanische Perspektive auf den ersten „allamerikanischen“, von Weißen wie Schwarzen bewunderten Sporthelden vor. Die Historikerin zeigt anhand von Presserecherchen, wie es in einer von Weißen dominierten Öffentlichkeit gelang, die Bedeutung der sensationellen Erfolge des Afro-Amerikaners Joe Louis so zu interpretieren, dass sie nicht als Bedrohung weißer Vorherrschaft wahrgenommen werden konnten, sondern als deren Bestätigung.

*Thomas Müller
Wolf-Dietrich Junghanns*

Gregor Ritschel, Thomas Müller

Big Data als Theorieersatz?

Wer eine Chiffre für unser digitales Zeitalter sucht und für die gemischten Gefühle, die sein Heraufziehen begleiten, kommt an „Big Data“ nicht vorbei. Ursprünglich gedacht als Bezeichnung für Datenmengen, die zu riesig und zu komplex sind, um sie auf herkömmliche Art und Weise zu verarbeiten, assoziiert man „Big Data“ mittlerweile mit der Datensammelwut von Internetkonzernen und Geheimdiensten sowie den daraus resultierenden Optionen der Analyse, Prognose und Kontrolle von Verhalten. Seit einigen Jahren lässt sich ein regelrechter Medienhype um „Big Data“ beobachten. Neonfarbene Nullen und Einsen rauschen, als Sinnbild allgegenwärtiger, aber schwer erfahrbarer Datenströme, über Zeitschriftencover oder Fernsehschirme. Jene Bilderwelt zeugt von Neugier, aber auch von einer gewissen Hilflosigkeit, Big Data zu begreifen. Dem exponentiellen Wachstum digitaler Daten ist kulturell kaum beizukommen, wengleich die Möglichkeiten der Datenaufbereitung zunehmen und eine Anwendungseuphorie vorherrscht, die kein gesellschaftliches Handlungsfeld vor Big Data verschont und auch in Alltag und Privatsphäre eingreift. „Data Scientist“ wird vermutlich bald einer der angesagtesten Berufe sein – schon 2012 galt er der *Harvard Business Review* als „Sexiest Job of the 21st Century“.

Wirtschaft, Politik, Alltagskommunikation, Öffentlichkeit, Kultur und Sport – all das wird von Big Data Schritt für Schritt durchdrungen. Doch was bedeutet Big Data für Wissenschaft und Forschung, vor allem: für die Sozial- und Geisteswissenschaften? Diese Frage steht im Mittelpunkt des vorliegenden Themenschwerpunkts. Sie zu stellen, liegt nicht zuletzt deshalb

nahe, weil sich mit dem Aufstieg von Big Data eine Provokation verbindet, die direkt auf das Selbstverständnis der Sozial- und Geisteswissenschaften zielt. Auf den Punkt gebracht hat sie der britische Journalist und Unternehmer Chris Anderson vor fast zehn Jahren, als er im Internet-Magazin *Wired* mit markigen Worten das Ende der Theorie und die Überflüssigkeit wissenschaftlicher Methoden verkündete: „Wir leben in einer Welt, in der riesige Datenmengen und angewandte Mathematik alle anderen Hilfsmittel ersetzen, die man sonst noch so anwenden könnte. Ob in der Linguistik oder in der Soziologie: Raus mit all den Theorien menschlichen Verhaltens. Vergessen Sie Taxonomien, die Ontologie und die Psychologie! Wer weiß schon, warum Menschen sich so verhalten, wie sie sich gerade verhalten? Der springende Punkt ist, dass sie sich so verhalten und dass wir ihr Verhalten mit einer nie gekannten Genauigkeit nachverfolgen und messen können. Hat man erst einmal genug Daten, sprechen die Zahlen für sich selbst“ (Anderson 2008/2013: 126).

Es wäre voreilig, Andersons Abgesang auf wissenschaftliche Theorien und Methoden als kenntnisfreie Spinnerei eines Laien abzutun und ihn, da der Autor sich inzwischen viel zurückhaltender äußert, zu den Akten zu legen. Trotz offensichtlicher Übertreibungen und Ungenauigkeiten steckt in Andersons Thesen eine radikale Vision, in der Erkenntnis aus der Datenfülle entspringt und uns die Umwege und Sackgassen der Reflexion erspart. Es ist der Traum von den nackten Zahlen, die für sich sprechen und keiner Interpretation bedürfen. In ihm haben weder Theorien noch Methoden

Platz. Gleichmaßen faszinierend wie verstörend ist die Vorstellung, dass sich dieser Traum mit Big Data erfüllen könnte.

Doch der Reihe nach. Andersons provokanten Thesen liegt die Beobachtung zugrunde, dass sich heute unvorstellbar große Datenmengen sammeln lassen, und zwar automatisch. Dies allein ist schon für empirische Forschung eine bemerkenswerte Entwicklung, die die bisherige Praxis des Datenerhebens verändert und Daten viel leichter verfügbar macht. Hinzu kommt, und das ist entscheidend, dass sich auch die Datenauswertung automatisiert und algorithmenbasiert vollzieht. Das Stichwort hierfür ist *Data-Mining* – das systematische Durchforsten von Massendaten nach interessanten Informationen und Zusammenhängen. An dieser Stelle setzen Andersons Überlegungen an: Wenn Daten automatisch, von Computern und Algorithmen, analysiert werden können, wozu müssen Forscher/-innen dann noch Hypothesen formulieren und prüfen? Digitale Datenanalyse braucht keine Hypothesen, denen theoretische Überlegungen zugrunde liegen. Digitale Datenanalyse benötigt strenggenommen auch keine Forscher/-innen mehr, denn den Sachverstand, der bei den verschiedenen Auswahlentscheidungen im Forschungsprozess gefragt ist, sollen auch Maschinen bzw. Algorithmen entwickeln können. In diesem Zusammenhang ist die Rede von *machine learning*. Hält man sich an eine Unterscheidung von Niklas Luhmann und begreift Big Data als sinnstrukturiertes Phänomen, so ließe sich zuspitzen: In *sachlicher* Hinsicht ersetzt Big Data Theorien, denn Daten und deren Verknüpfung reichen aus, um Modelle zu entwickeln und neue Einsichten zu gewinnen. In *zeitlicher* Hinsicht beschleunigen sich Datenerfassung und -auswertung dank Big Data massiv, sie laufen in Echtzeit ab, also simultan zur Datenerzeugung. In *sozialer* Hinsicht treten mit Big Data Computer und Algorithmen als lernende Maschinen an die Stelle von Personen, die auf der Grundlage von Erfahrung und Expertise urteilen.

Auch jene Vordenker und Experten, denen Andersons Verabschiedung der Theorie zu weit gehen dürfte, nehmen an, dass sich die Wissenschaften im Angesicht von Big Data

grundlegend verändern werden. Der österreichische Jurist Viktor Mayer-Schönberger erwartet beispielsweise, dass die Sozialwissenschaften „sozialwissenschaftlicher“ werden, weil sie mit belastbareren Ergebnissen als bislang aufwarten können. Auch die Geisteswissenschaften könnten von Big Data profitieren und „durch Quantifikation und Datenanalyse neue Erkenntnisse gewinnen“. Auch wenn die riesigen Datenmengen (nur) als „Rohstoff für den Erkenntnisprozess“ genutzt würden, bietet Big Data doch einen „neuen Zugang zur Welt“, eine „Infrastruktur der Aufklärung für das 21. Jahrhundert“ (Mayer-Schönberger 2016).

Weniger zuversichtlich sind hingegen einige Lebenswissenschaftler, die befürchten, ein „Daten-Tsunami“ (Müller-Jung 2013) könne die Forschung überrollen und kollabieren lassen. Big Data wäre dieser Lesart zufolge weniger ein neuer vielversprechender Zugang zur Welt als vielmehr ein Beitrag zur Verbreitung des Messie-Syndroms in den Wissenschaften: Daten werden um ihrer selbst willen gehortet, ohne dass man eine Fragestellung hat oder etwas „auf die Spur kommen möchte“, wie der Wissenschaftsforscher Hans-Jörg Rheinberger kritisiert. Unordnung und digitaler Kollaps seien die sich abzeichnenden Gefahren, wenn die wissenschaftliche Analyse mit den stetig wachsenden Datenbergen nicht mehr Schritt halten kann. Während in früheren Zeiten das Problem des Datenmangels bestand, bestehe heute eher Analysemangel, so der Neurobiologe Hans Hofmann (vgl. ebd.). Die epistemologische Revolution, die der Übergang von der hypothesengeleiteten zur datengeleiteten Forschung darstellen soll, vollzieht sich offenbar nicht so unkompliziert und reibungslos, wie seine Apologeten wünschen (vgl. Rheinberger 2007).

Das Ende des Samplings

Doch bleiben wir bei den Sozialwissenschaften und dem technologisch bedingten Wandel, der sich dort gegenwärtig vollzieht. Viktor Mayer-Schönberger und Kenneth Cukier verdeutlichen, inwiefern die traditionelle empirische Wirtschafts- und Sozialforschung heute unter Druck gerät: „Seit dem 19. Jahrhundert ist die

Gesellschaft auf Stichproben (sog. Samples) angewiesen, wenn es um sehr große Zahlen geht. Dieser Zwang zu repräsentativen Stichproben ist ein Relikt aus einer Zeit, als Information etwas Außergewöhnliches war, ein Ergebnis der Grenzen der Informationsverarbeitung im analogen Zeitalter. Vor der Einführung leistungsfähiger Digitaltechnologien erkannten wir die Samples nicht als Behinderung, sondern nahmen sie als unvermeidlich hin. Aber indem wir alle Daten sammeln, sind wir nun in der Lage, Einzelheiten zu sehen, die vorher nicht erkennbar waren, solange wir nur kleine Datenmengen verarbeiten konnten. Big Data gibt uns einen besonders deutlichen Einblick in Details: Unterkategorien und Marktnischen etwa, die der Stichprobenanalyse unzugänglich blieben“ (Mayer-Schönberger/Cukier 2013: 20f.). Die mögliche Unschärfe der einzelnen Datenmessung ist hier kein Problem mehr, insofern der Mittelwert fast unverändert bleibt (ebd.: 21). Zugleich ist die „hohe Auflösung“ des mit Big Data generierten Bildes als ein weiterer Vorteil gegenüber dem Sampling zu betrachten, da man nun tiefer in die Daten „hineinzoomen“ kann. Durch die neuen Sortierungsmöglichkeiten können beispielsweise verschiedenste Personengruppen aus der Betrachtung herausgenommen werden. Mayer-Schönberger und Cukier zufolge führt dieser methodische Wandel zur „Abwendung von der jahrtausendealten Suche nach kausalen Zusammenhängen“ (ebd.: 22). „Als Menschen sind wir darauf ausgelegt, bei allem nach seiner Ursache zu fragen, auch wenn das oft schwierig ist und uns vielleicht auf eine falsche Fährte führt. In der Big-Data-Welt dagegen müssen wir uns nicht auf Kausalitäten festlegen, sondern können viel öfter auch nach Mustern und Korrelationen in den vorliegenden Daten Ausschau halten, die uns neuartige und wertvolle Erkenntnisse gewähren. Die Korrelationen sagen uns nicht, warum etwas geschieht, aber sie machen uns darauf aufmerksam, dass etwas geschieht“ (ebd.).

In Sozialwissenschaften, die auf Big Data zurückgreifen, steht nicht länger eine Hypothese am Beginn der Forschung, sondern ein exploratives Vorgehen. Das deduktiv-nomologische Modell wissenschaftlicher Forschung wird

ersetzt durch eine theoriefreie Suche nach Korrelationen, die Hinweise auf Phänomene gibt, die erst im Anschluss weiter erforscht werden (Röhle 2014: 167). Wenn man die Grenzen zwischen der traditionellen Logik der Forschung und Big Data nicht ganz so scharf zieht, ließe sich aber auch fragen, ob jene großen Datensätze als *Vorstufe* des bisherigen Erkenntnisverfahrens zu begreifen sind, indem also etwa Korrelationen festgestellt werden, die erst im Anschluss durch die eigentliche sozialwissenschaftliche Forschung auf Kausalitätsbeziehungen hin überprüft werden (vgl. Mayerl 2015)? Schließlich steht am Anfang der Hypothesenbildung immer auch ein erfahrungsbasiertes „Verdachtsmoment“. Fraglich ist jedoch, inwiefern die Technologie des maschinellen Lernens eigenständige Hypothesen zu bilden vermag, die in den Sozialwissenschaften Anwendung finden können.

Das Versprechen von „Big Data“ steckt in den „drei Vs“: Mit *Volume* ist die Datenmenge gemeint, mit *Velocity* die hohe Geschwindigkeit der Datenverarbeitung und mit *Variety* die Datenvielfalt, d. h. die Möglichkeit des Rückgriffs auf unterschiedliche Arten von Datensätzen. Insbesondere die Echtzeitverarbeitung und -darstellung von Stimmungen und Trends im „Gefühlshaushalt“ einer Bevölkerung tragen zur Faszination von Big Data bei. Schon im US-Präsidentenwahlkampf 2012 stützte sich die Kampagne Barack Obamas auf Big Data-Analysen, die halfen, die stets begrenzten Mittel möglichst zielgenau einzusetzen (Moorstedt 2013). Auch zum Wahlsieg Donald Trumps soll Big Data beigetragen haben: Die Datenfirma *Cambridge Analytica* stellte Trumps Kampagne ihr gesammeltes Wissen über angeblich mehr als 230 Millionen erwachsene Amerikaner zur Verfügung und ermöglichte so eine Ansprache potentieller Wähler/-innen in den sozialen Netzwerken, die deren individuellen Präferenzen gleicht (vgl. Müller von Blumencron 2016). Ob dies die Präsidentenwahl entschieden hat, ist freilich umstritten. In Anbetracht der Umfragen, die einen Wahlerfolg Hillary Clintons vorhersagten und damit völlig danebenlagen, sind leise Zweifel an den prognostischen Potentialen von Big Data vielleicht nicht unangebracht.

Daten: roh oder gekocht?

Während man in England schon im 17. Jahrhundert von „data“ sprach (Rosenberg 2014), bezeichnete man im 20. Jahrhundert große Datensätze bzw. Massendaten (z. B. Ergebnisse von Volkszählungen), zu deren Auswertung man Supercomputer benötigte, als „Big Data“. Mittlerweile ist hieraus ein unscharfer Begriff geworden, der nicht zuletzt für eine Erkenntnisweise bzw. Analysemethode steht, die darauf gerichtet ist, in digitalisierten Datenbeständen Muster und Korrelationen mit Hilfe von Algorithmen zu erkennen, die ihrerseits einer hohen Rechenleistung bedürfen. Dies verweist auf die technischen und technologischen Dimensionen von Big Data. Nicht zu übersehen ist aber auch die mythologische Dimension: Big Data-Technologie wird oftmals mit dem Glauben verknüpft, große Datensätze führten unmittelbar zu mehr Objektivität oder zu einer höheren Form der Wahrheit (boyd/Crawford 2013: 188f.). Eine wissenschaftshistorische Perspektive zeigt, dass in den Geistes- und Sozialwissenschaften immer wieder der Wunsch auftauchte, durch eine Quantifizierung der Forschung die menschliche Insuffizienz zu überwinden (Röhle 2014: 163). In jüngster Zeit äußerte sich dieser Wunsch etwa in Plädoyers für eine neue, auf Big Data gestützte Sozialwissenschaft als Teilbereich der „Digital Humanities“. Daneben finden sich aber auch zahlreiche skeptische Stimmen. Der Literatur- und Medienwissenschaftler Roberto Simanowski gibt folgendes zu bedenken: „Aus Sicht einer geisteswissenschaftlichen Soziologie, wie sie Adorno der empirisch-positivistischen Soziologie entgegenstellt, ist der Algorithmus nicht Liebesobjekt, sondern Feindbild: weil er das Gesellschaftliche auf Mathematik reduziert und abweichende, alternative Positionierungen blockiert“ (Simanowski 2015: 85).

Angesichts der heute wieder diskutierten Frage, ob Daten als „roh“ oder als „gemacht“ zu begreifen sind (Gitelman 2013), scheint die Einsicht, dass Daten stets theoriegeleitet gewonnen werden, in Vergessenheit zu geraten. Daten lassen sich nicht einfach aus Natur oder Kultur herausnehmen, vielmehr müssen sie in den Naturwissenschaften wie in den Geistes- und

den Sozialwissenschaften durch eine komplexe Kette von Übersetzungsleistungen erst erzeugt bzw. gemacht werden. Dieses „doing data“ zeigt der Wissen(schaft)ssoziologe Bruno Latour an einem Beispiel: Forscher, die die Versandung des Regenwaldes untersuchen, entnehmen Bodenproben. Von der Bodenprobe wird über den Probenkasten hin zur tabellarischen Darstellung vorangeschritten. In diesem Prozess zeigt sich die von Menschen gemachte Generierung einer Referenzkette (Latour 2000: 36ff.). Die Tatsache der Herstellung von Daten ändert aber nichts an deren Objektivität. Denn die Kette der Übersetzungen von der Grafik bis zurück in den Boden des Amazonas-Regenwalds lässt sich jederzeit rückwärts abschreiten. Auch „rohe“ Daten, wie die Daten der tabellarischen Visualisierungen der Versandung, sind demnach gemachte Übersetzungsschritte. Im Unterschied zu den Bodenproben, auf die Latour verweist, sind digitale Daten bzw. digitale Spuren in sozialen Netzwerken, die eben noch nicht aufbereitet wurden, weit weniger leicht auf Tatsächlichkeiten rückführbar und daher oftmals von zweifelhafter Qualität (Venturini u. a. 2015: 17ff.). Aus Latours Beispiel lässt sich aber auch schließen, dass selbstlernende Algorithmen, die angesichts von festgestellten Korrelationen Hypothesen bilden, in ihren Schritten der Forschungscommunity transparent sein müssen, wenn sie sinnvoll in die Wissenschaft eingebunden werden sollen.

Die Möglichkeiten von Big Data suggerieren: Das Zeitalter des Verstehens ist vorbei, Theorien und Modelle sind obsolet, man braucht nur noch Korrelationen zu betrachten. Dem Anschein nach lässt sich mit den Ergebnissen besser und bequemer leben. Angesichts dieser Substitution aber fragt man sich: Sind die Algorithmen unsere nützlichen Prothesen angesichts postmoderner Unübersichtlichkeiten? Beobachten wir nur eine weitere Etappe technischen Könnens, die uns von notwendiger Arbeit entlastet und uns mehr Raum zur Entfaltung bietet? Oder stehen wir am Ende der kritischen Tradition der Aufklärung, die sich an normativen Idealen orientierte? Geht mit dem Austausch der Erkenntnismittel auch eine Veränderung des Denkens selbst einher,

wie manche meinen (vgl. boyd/Crawford 2013: 192)?

Macht und Mittel

Einst weckte das Internet die Hoffnung auf das Ende aller Monopole und auf eine freie und rationale Kommunikation. Heute dominieren Daten sammelnde und zu kommerziellen Zwecken verwertende Big Player wie Google, Facebook oder Amazon die Netzlandschaft. Lev Manovich (2014: 77) weist auf die gesellschaftliche Spaltung zwischen der Mehrheit der Datenproduzenten, den wenigen Datensammlern und der verschwindend kleinen Gruppe der Datenanalysten hin. Er gibt zu bedenken, dass sich mit Hilfe von Big Data potenziell auch interessante soziologische Fragen klären lassen, es aber viel öfter darum gehe, die Antworten auf die Frage, was Menschen zukünftig wollen, effektiv zu kapitalisieren (ebd.). In eine ähnliche Richtung argumentiert auch Google News-Chef Richard Gingras, der sich mit folgender Bemerkung über wertvolle Inhalte, für die Nutzer auch zu zahlen bereit sind, zitieren lässt: „Der wertvollste Inhalt ist jener, der uns interessiert, weil er weit über das hinausgeht, für das wir glaubten, uns zu interessieren.“ (Müller-Wirth/Wefing 2016).

Neben den Chancen, die Big Data bietet, droht auch in den Wissenschaften ein „New Digital Divide“. Der Zugriff auf und die Nutzung von Big Data könnten in der Forschungslandschaft Brüche erzeugen, die zu Hierarchien anwachsen und Monopolbildungen begünstigen. Auch in den Geistes- und Sozialwissenschaften entfalten sich „Wahrheiten“ nicht von allein, sondern Wissenschaftler/-innen setzen sich und ihre Projekte und Perspektiven in dem Maße durch, wie es ihnen gelingt, ihre Netzwerke, Ressourcen und ihr symbolisches Kapital zu mobilisieren. Big Data könnte sich insofern auch in den Sozialwissenschaften als Machtfaktor erweisen. Umgekehrt muss zukünftig auch die oftmals undurchsichtige Entscheidungsmacht algorithmischer Prozesse auf die Gesamtgesellschaft kritisch in den Blick genommen werden. Immer mehr alltägliche Entscheidungen werden in der Welt der Massendaten und des Internets

der Dinge an Computer und Algorithmen überantwortet. Vermeintlich rein technische Entscheidungsprozesse können aber durchaus politisch wirken.

Zu den Beiträgen

Der Themenschwerpunkt beleuchtet die Rolle, die Big Data in den Sozial- und Geisteswissenschaften spielt, aus unterschiedlichen Perspektiven. So entsteht ein gleichermaßen facettenreiches wie differenziertes Bild, das Fragen von wissenschaftstheoretischer und disziplinpolitischer Relevanz ebenso berührt wie Fragen, die Datenschutz, politischen Aktivismus und Technikgläubigkeit betreffen. Werden Theorien zukünftig obsolet? Gibt es einen Paradigmenwechsel in den Sozialwissenschaften von der hypothesen- zur daten- geleiteten Wissenschaft? Vollzieht sich ein programmatischer Wandel von der deduktiv-nomologischen hin zur induktiv-explorativen Logik der Forschung? Machen Korrelationen die Suche nach Kausalitäten überflüssig? Führt die Fixierung auf die Suche nach Korrelationen zu falschen Schlüssen? Wie könnte eine Kombination von klassischer Sozialwissenschaft und digitalen Technologien aussehen? Welche institutionellen Entwicklungen im Feld der Sozial- und Geisteswissenschaften zeichnen sich ab und wie sind diese mit Machtfragen bzw. Machtkonstellationen innerhalb der Sozialwissenschaften verknüpft? Ist die Überantwortung an Algorithmen problematisch, da diese eine Hörigkeit gegenüber dem Vergangenen in sich tragen? Und was bedeutet dies für die angewandten Sozialwissenschaften? Wenn Algorithmen optimieren sollen: Was ist die Norm? Und schließlich: Sind Daten „roh“ oder „gemacht“? Welcher ontologische Status ist ihnen zuzuordnen? Und was lässt sich über unsere zukünftige Epistemologie sagen?

Jan-Felix Schrape unterscheidet vier Stufen bzw. Phasen der kommunikations- und informationstechnischen Durchdringung der Gesellschaft. Im Zentrum seines Beitrags stehen die Hoffnungen und Ängste, die mit den jeweiligen technologischen Neuerungen verbunden wurden. Überblickt man die historische

Spanne vom Entstehen der Informationsgesellschaft in den 1960er-Jahren bis hin zur Big Data-Gesellschaft, die seit den 2010-Jahren entsteht, so zeigt sich, dass die konzeptualisierende Gesellschaftswissenschaft nicht nur Erklärungen für gesellschaftliche Wandlungen bieten kann, sondern auch ein Mittel darstellt, um die periodisch wiederkehrende Technikgläubigkeit einer breiten Öffentlichkeit kritisch zu diskutieren und zu entzaubern. Aus einer datenhistorischen Metaperspektive geht es nicht darum, die Erwartung bzw. Befürchtung, Big Data könne Theorie ersetzen, als richtig oder falsch zu bewerten. Vielmehr wird die Genese jener Annahmen und der aus ihnen resultierenden Konsequenzen für die Gesellschaftswissenschaft ins Zentrum der Aufmerksamkeit gerückt.

Big Data-Technologien erheben Daten und kombinieren diese auf der Suche nach Korrelationen. Dies ist jedoch nicht Wissenschaft im eigentlichen Sinne. In einem philosophiegeschichtlichen Abriss betont *Klaus Mainzer*, dass am Anfang allen Wissens die Frage nach dem Warum stand. Diese Frage sei bis heute relevant geblieben. So nütze die statistische Vorausberechnung von Kriminalität wenig, wenn die sozialen und politischen Ursachen des Phänomens nicht begriffen werden. Mainzer zeigt, dass seit jeher Empiriker, Skeptiker und Philosophen über Nutzen und Notwendigkeit von Theorien gestritten haben. Doch gerade heute erweist sich, dass auch Big Data nicht ohne jenes logisch-mathematische Denken möglich ist, das am Anfang der Philosophie stand. Die immer komplexer werdende Welt, auch die „Soziodynamik“ mitsamt ihren menschlichen Faktoren, ist für Mainzer prinzipiell berechenbar, doch nicht ohne Theorien, Forschungshypothesen, Experimente und Falsifikationen.

Christian Wadephul setzt sich mit Big Data-Analysen als Grundlage maschinellen Lernens auseinander. Sogenannte maschinell lernende Artefakte (MLA) lernen abduktiv. Die Abduktion bildet angesichts neuer Ereignisse Hypothesen über die Funktionsweise der Welt. Sie steht somit jenseits von Induktion und Deduktion. Abduktionen hervorbringende maschinelle Lernsysteme erstellen Theorien auf Grundlage ihres datengetriebenen Lernens,

die dazu dienen können, große Datenmengen komprimiert darzustellen und zu deuten. Diese von der KI erstellten Abduktionen können jedoch auch falsch sein, insofern Korrelationen und Muster erkannt werden, die entweder falsch oder zufällig sind. Ganz gleich, ob MLAs in der Steuerung von selbstfahrenden Autos oder als nützliches Instrument bei der Durchforstung von digitalisierten Wissensbeständen eingesetzt werden – der Mensch muss sie auch zukünftig mit Umsicht überwachen.

Die aktuelle rechtswissenschaftliche Diskussion über Big Data greifen *Roland Broemel* und *Hans-Heinrich Trute* in ihrem Beitrag kritisch auf. Sie wenden sich vor allem gegen die strukturelle Engführung auf die datenschutzrechtliche Perspektive und die damit einhergehende Neigung, Daten als Repräsentation der Wirklichkeit zu deuten und jegliche sozial-konstruktive Dimension auszublenden. Die von den Autoren vorgeschlagene Kategorisierung von Daten legt die Einsicht nahe, dass die datenschutzrechtliche Perspektive sich nur auf direkt personenbezogene Daten beziehen lässt. Die für Big Data-Anwendungen eigentlich relevanten (oft anonymisierten) Massenbeobachtungsdaten, die von diesen abgeleiteten Daten und weiterhin generierte Meta-Daten fallen hierbei jedoch unter den Tisch. Wenn mittels Big Data anonymisierte Clusterdaten generiert werden, dann erzeugen diese in der Folge Zuschreibungen gegenüber Einzelnen, die etwa bei der Gewährung eines Kredites problematisch werden können. Die Frage, wie mit solchen Big Data-Anwendungen rechtstheoretisch umgegangen werden sollte, ist bisher offen. Der Ruf nach mehr Datenschutz geht hierbei jedenfalls ins Leere.

Stefan Böschen, *Georg Huber* und *René König* verweisen auf die politische Dimension von Big Data. Sie fragen nach einem Ansatzpunkt, um Big Data-Technologie und deren Folgen aus gesellschaftswissenschaftlicher Sicht zu verstehen. Die soziale Wirksamkeit von Algorithmen wird dabei in den Vordergrund gestellt. Es gelte, die Governance mit und durch Algorithmen zu betrachten. Um diese Wirksamkeit aufzuschließen, greifen die Autoren auf Ulrich Becks Konzept der Subpolitik zurück, das das politische Wirken

der „Neuen Sozialen Bewegungen“ der 1980er Jahre zu erfassen suchte. Die heutigen Big Data-Algorithmen deuten die Autoren als eine Form passiv-struktureller Subpolitik. Als solche sind mit Big Data operierende Algorithmen eine Form von Politik, die sich selbst nicht als politisch begreift, wohl aber gesellschaftliche Verbindlichkeiten schafft. Weder sind Plattformen des Cyberaktivismus wie Avaaz neutrale Plätze der politischen Selbstorganisation noch präsentieren Online-Suchmaschinen wie Google im Hinblick auf politische Inhalte neutrale Trefferlisten. Big Data macht, so das Argument des Textes, Politik, ohne sich als politisch auszuweisen.

Abschließend formulieren *Jochen Mayerl* und *Katharina Anna Zweig* drei Thesen zur Zukunft der Soziologie in der digitalen Gesellschaft. Sie wenden sich damit gegen ein selbstverschuldetes Abdanken der Soziologie als Leitdisziplin für das Erklären und Verstehen des Sozialen. Soziologische Datenerhebungen sind bislang nicht einmal ansatzweise als Big Data zu qualifizieren. Big Data mag aus soziologischer Sicht unsauber sein, doch sind die soziologischen Samples oft ebenfalls nicht frei von Mängeln. Eine Berechnung der Gesellschaft auf Grundlage der so oder so anfallenden Datenmassen kann potentiell auch von Informatikern geleistet werden. Dabei ginge aber das theoretische Know-how der Soziologie verloren. Mayerl und Zweig halten die Entwicklung interdisziplinärer Forschungsprojekte eindeutig für den besseren Weg. Nur so könnten die neuen technologischen Möglichkeiten von Big Data mit dem theoretischen Erfahrungsschatz der Soziologie sinnvoll kombiniert werden.

Literatur

- Anderson, Chris (2008/2013): Das Ende der Theorie. Die Datenschwemme macht wissenschaftliche Methoden obsolet. In: Geiselberger, Heinrich; Moorstedt, Tobias (Hrsg.): Big Data. Das neue Versprechen der Allwissenheit. Berlin: Suhrkamp, S. 124-130.
- boyd, danah; Crawford, Kate (2013): Big Data als kulturelles, technologisches und wissenschaftliches Phänomen. Sechs Provokationen. In: Geiselberger, Heinrich; Moorstedt, Tobias (Hrsg.): Big Data. Das neue Versprechen der Allwissenheit. Berlin: Suhrkamp, S. 187-218.
- Davenport, Thomas H.; Patil, D. J. (2012): Data Scientist: The Sexiest Job of the 21st Century. In: Harvard Business Review, October 2012. URL: <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century/> (Stand: 10.12.2016).
- Gitelman, Lisa (Hrsg.) (2013): Raw Data is an Oxymoron. Cambridge: MIT Press.
- Latour, Bruno (2000): Die Hoffnung der Pandora. Untersuchungen zur Wirklichkeit der Wissenschaft. Frankfurt am Main: Suhrkamp.
- Manovich, Lev (2014): Trending. Verheißungen und Herausforderungen der Big Social Data. In: Reichert, Ramón (Hrsg.): Big Data. Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie. Bielefeld: Transcript, S. 65-83.
- Mayer-Schönberger, Viktor (2016): Ein Instrument des Verstehens. Interview: Stefan Schmitt. In: Die Zeit, 6. Oktober 2016, Nr. 42, S. 37. URL: <http://www.zeit.de/2016/42/big-data-wissenschaft-folgen-viktor-mayer-schoenberger/komplettansicht> (Stand: 10.12.2016).
- Mayer-Schönberger, Viktor; Cukier, Kenneth (2013): Big Data: Die Revolution, die unser Leben verändern wird. München: Redline.
- Mayerl, Jochen (2015): Bedeutet 'Big Data' das Ende der sozialwissenschaftlichen Methodenforschung? In: Soziopolis.de, 29.12.2015. URL: <http://www.sozio.polis.de/beobachten/wissenschaft/artikel/bedeutet-big-data-das-ende-der-sozialwissenschaftlichen-methodenforschung/> (Stand: 10.12.2016).
- Moorstedt, Tobias (2013): Obamas Datenakrobaten. In: Geiselberger, Heinrich; Moorstedt, Tobias (Hrsg.): Big Data. Das neue Versprechen der Allwissenheit. Berlin: Suhrkamp, S. 35-54.
- Müller von Blumencron, Mathias (2016): Im neuen Propagandazeitalter. Donald Trump gewann mit einer Kampagne kunstvoller digitaler Manipulation. In: Frankfurter Allgemeine Zeitung, 6. Dezember 2016, Nr. 285, S. 8.
- Müller-Jung, Joachim (2013): Wird Big Data zur Chiffre für den digitalen GAU? In: Frankfurter Allgemeine Zeitung, 6. März 2013, Nr. 55, Seite N1.
- Müller-Wirth, Moritz; Wefing, Heinrich (2016): Ich. Google, Facebook und Co. wollen die Nachrichten revolutionieren. In: Die Zeit, 3. November 2016, Nr. 46, S. 4. URL: <http://www.zeit.de/2016/46/journalismus-facebook-twitter-silicon-valley-demokratie-politik/komplettansicht> (Stand: 10.12.2016).
- Rheinberger, Hans-Jörg (2007): Wie werden aus Spuren Daten, und wie verhalten sich Daten zu Fakten? In: Gugerli, David u. a. (Hrsg.): Nach Feierabend. Zürcher Jahrbuch für Wissenschaftsgeschichte Nr. 3: Datenbanken. Zürich, Berlin: Diaphanes, S. 117-125.

- Röhle, Theo (2014): Big data – Big Humanities? Eine historische Perspektive. In: Reichert, Ramón (Hrsg.): Big Data. Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie. Bielefeld: Transcript, S. 157-172.
- Rosenberg, Daniel (2014): Daten vor Fakten. In: Reichert, Ramón (Hrsg.): Big Data. Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie. Bielefeld: Transcript, S. 133-156.
- Simanowski, Roberto (2014): Data Love. Berlin: Matthes & Seitz.
- Venturini, Tommaso; Latour, Bruno; Meunier (2015): Eine unerwartete Reise. Einige Lehren über die Kontinuität aus den Erfahrungen des Sciences Po médialab. In: Süssenguth, Florian (Hrsg.): Die Gesellschaft der Daten. Über die digitale Transformation der sozialen Ordnung. Bielefeld: Transcript, S. 17-39.

Berliner Debatte Initial 27 (2016) 4

Sozial- und geisteswissenschaftliches Journal

© **Berliner Debatte Initial** e.V.,
Ehrenpräsident Peter Ruben.
Berliner Debatte Initial erscheint viermal jährlich.

Redaktionsrat: Harald Bluhm,
Wladislaw Hedeler, Cathleen Kantner,
Rainer Land, Udo Tietz, Andreas Willisch.

Redaktion: Ulrich Busch, Erhard Crome, Wolf-
Dietrich Junghanns, Raj Kollmorgen, Thomas
Möbius, Robert Stock, Matthias Weinhold,
Johanna Wischner.

Redaktionelle Mitarbeit: Gregor Ritschel, Ben-
jamin Sonntag.

Verantwortl. Redakteur: Thomas Müller. **Vi.S.P.**
für dieses Heft: Thomas Müller

Satz: Rainer Land.

Copyright für einzelne Beiträge ist bei der Redak-
tion zu erfragen.

E-Mail: redaktion@berlinerdebatte.de
<http://www.redaktion.berlinerdebatte.de/>

Berliner Debatte Initial erscheint bei
WeltTrends, Medienhaus Babelsberg
August-Bebel-Straße 26-53
D-14482 Potsdam
www.welttrends.de

Preise: Einzelheft 15 €,
Jahresabonnement 40 €, Institutionen 45 €,
Studenten, Rentner und Arbeitslose 25 €.
Für ermäßigte Abos bitte einen Nachweis (Kopie)
beilegen. Das Abonnement gilt jeweils für ein
Jahr und verlängert sich um jeweils ein Jahr, wenn
nicht sechs Wochen vor Ablauf gekündigt wird.

Bestellungen: Einzelhefte im Buchhandel;
Einzelhefte (gedruckt oder als PDF) und einzelne
Artikel (als PDF) im Webshop:
<http://shop.welttrends.de/>
oder per E-Mail:
bestellung@welttrends.de
oder telefonisch: +49/331/721 20 35
(Büro WeltTrends)

Abonnement per Mail, telefonisch oder per Post

bestellung@welttrends.de
+49/331/721 20 35

WeltTrends, Medienhaus Babelsberg
August-Bebel-Straße 26-53
D-14482 Potsdam

Ich bestelle ein Abonnement der Berliner Debatte Initial ab Heft

- Jahresabonnement 40 € (Institutionen 45 €).
- Abonnement ermäßigt 25 Euro (Studenten, Rentner, Arbeitslose etc.),
Nachweis bitte beilegen.

Name:

Straße, Nr.:

Postleitzahl: Ort: Telefon:

Ich weiß, dass ich diese Bestellung innerhalb von 10 Tagen bei der Bestelladresse schriftlich wider-
rufen kann.

Datum: Unterschrift: